



OIPE

## RAW SEQUENCE LISTING

DATE: 09/09/2002

PATENT APPLICATION: US/10/050,216B

TIME: 12:50:08

Input Set : A:\CORRECTEDSEQUENCE LISTING.txt

Output Set: N:\CRF3\09092002\J050216B.raw

4 <110> APPLICANT: Curtis, Rory A.J., Lora, Jose M.  
6 <120> TITLE OF INVENTION: 46798, A Human Matrix Metalloprotease and  
7 Uses Therefore  
9 <130> FILE REFERENCE: MPI2001-014P1RNM  
11 <140> CURRENT APPLICATION NUMBER: 10/050,216B  
C--> 13 <141> CURRENT FILING DATE: 2002-08-26  
15 <150> PRIOR APPLICATION NUMBER: 60/262,252  
17 <151> PRIOR FILING DATE: 2001-01-16  
19 <160> NUMBER OF SEQ ID NOS: 10  
21 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
23 <210> SEQ ID NO: 1  
24 <211> LENGTH: 2310  
25 <212> TYPE: DNA  
26 <213> ORGANISM: Homo sapiens  
28 <220> FEATURE:  
29 <221> NAME/KEY: CDS  
30 <222> LOCATION: (317)...(1651)  
32 <400> SEQUENCE: 1  
33 gtcgacccac gcgtccggcc gggcctccgc cccctccgcc tgcctttcct tcctccctcc 60  
34 ctcggtcccc ggggccggcg gacccgcggg caggcactgc ccgggctgga cgacgtctgg 120  
35 ccggtctcccg gcgaagggca gcggaggagc ggcccagagc gcgcagctag ggcactggcg 180  
36 aaaccccggg acagtccctc tccgtgcggg ggcgggcagc agcagtccca tccccggggt 240  
37 cccgggcgcg gctgactgcc ggctggttcc ctgcgcgcag tagctccccg agccgggctg 300  
38 caccggaggc ggcgag atg gtc gcg cgc gtc ggc ctc ctg ctg cgc gcc ctg 352  
39 Met Val Ala Arg Val Gly Leu Leu Leu Ala Leu  
40 1 5 10  
42 cag ctg cta ctg tgg ggc cac ctg gac gcc cag ccc gcg gag cgc gga 400  
43 Gln Leu Leu Leu Trp Gly His Leu Asp Ala Gln Pro Ala Glu Arg Gly  
44 15 20 25  
46 ggc cag gag ctg cgc aag gag gcg gag gca ttc cta gag aag tac gga 448  
47 Gly Gln Glu Leu Arg Lys Glu Ala Glu Ala Phe Leu Glu Lys Tyr Gly  
48 30 35 40  
50 tac ctc aat gaa cag gtc ccc aaa gct ccc acc tcc act cga ttc agc 496  
51 Tyr Leu Asn Glu Gln Val Pro Lys Ala Pro Thr Ser Thr Arg Phe Ser  
52 45 50 55 60  
54 gat gcc atc aga gcg ttt cag tgg gtg tcc cag cta cct gtc agc ggc 544  
55 Asp Ala Ile Arg Ala Phe Gln Trp Val Ser Gln Leu Pro Val Ser Gly  
56 65 70 75  
58 gtg ttg gac cgc gcc acc ctg cgc cag atg act cgt ccc cgc tgc ggg 592  
59 Val Leu Asp Arg Ala Thr Leu Arg Gln Met Thr Arg Pro Arg Cys Gly  
60 80 85 90  
62 gtt aca gat acc aac agt tat gcg gcc tgg gct gag agg atc agt gac 640  
63 Val Thr Asp Thr Asn Ser Tyr Ala Ala Trp Ala Glu Arg Ile Ser Asp

## RAW SEQUENCE LISTING

DATE: 09/09/2002

PATENT APPLICATION: US/10/050,216B

TIME: 12:50:08

Input Set : A:\CORRECTEDSEQUENCE LISTING.txt

Output Set: N:\CRF3\09092002\J050216B.raw

```

64          95          100          105
66 ttg ttt gct aga cac cgg acc aaa atg agg cgt aag aaa cgc ttt gca 688
67 Leu Phe Ala Arg His Arg Thr Lys Met Arg Arg Lys Lys Arg Phe Ala
68      110          115          120
70 aag caa ggg ggc gcc ctg gcg cac gcc ttc ctg ccc cgc cgc ggc gaa 736
71 Lys Gln Gly Gly Ala Leu Ala His Ala Phe Leu Pro Arg Arg Gly Glu
72 125          130          135          140
74 gcg cac ttc gac caa gat gag cgc tgg tcc ctg agc cgc cgc cgc ggc 784
75 Ala His Phe Asp Gln Asp Glu Arg Trp Ser Leu Ser Arg Arg Arg Gly
76          145          150          155
78 cgc aac ctg ttc gtg gtg ctg gcg cac gag atc ggt cac acg ctt ggc 832
79 Arg Asn Leu Phe Val Val Leu Ala His Glu Ile Gly His Thr Leu Gly
80      160          165          170
82 ctc acc cac tcg ccc gcg ccg cgc gcg ctc atg gcg ccc tac tac aag 880
83 Leu Thr His Ser Pro Ala Pro Arg Ala Leu Met Ala Pro Tyr Tyr Lys
84      175          180          185
86 agg ctg ggc cgc gac gcg ctg ctc agc tgg gac gac gtg ctg gcc gtg 928
87 Arg Leu Gly Arg Asp Ala Leu Ser Trp Asp Asp Val Leu Ala Val
88      190          195          200
90 cag agc ctg tat ggg aag ccc cta ggg ggc tca gtg gcc gtc cag ctc 976
91 Gln Ser Leu Tyr Gly Lys Pro Leu Gly Gly Ser Val Ala Val Gln Leu
92 205          210          215          220
94 cca gga aag ctg ttc act gac ttt gag acc tgg gac tcc tac agc ccc 1024
95 Pro Gly Lys Leu Phe Thr Asp Phe Glu Thr Trp Asp Ser Tyr Ser Pro
96          225          230          235
98 caa gga agg cgc cct gaa acg cag ggc cct aaa tac tgc cac tct tcc 1072
99 Gln Gly Arg Arg Pro Glu Thr Gln Gly Pro Lys Tyr Cys His Ser Ser
100      240          245          250
102 ttc gat gcc atc act gta gac agg caa cag caa ctg tac att ttt aaa 1120
103 Phe Asp Ala Ile Thr Val Asp Arg Gln Gln Gln Leu Tyr Ile Phe Lys
104      255          260          265
106 ggg agc cat ttc tgg gag gtg gca gct gat ggc aac gtc tca gag ccc 1168
107 Gly Ser His Phe Trp Glu Val Ala Ala Asp Gly Asn Val Ser Glu Pro
108      270          275          280
110 cgt cca ctg cag gaa aga tgg gtc ggg ctg ccc ccc aac att gag gct 1216
111 Arg Pro Leu Gln Glu Arg Trp Val Gly Leu Pro Pro Asn Ile Glu Ala
112 285          290          295          300
114 gcg gca gtg tca ttg aat gat gga gat ttc tac ttc ttc aaa ggc ggt 1264
115 Ala Ala Val Ser Leu Asn Asp Gly Asp Phe Tyr Phe Phe Lys Gly Gly
116          305          310          315
118 cga tgc tgg agg ttc cgg ggc ccc aag cca gtg tgg ggt ctc cca cag 1312
119 Arg Cys Trp Arg Phe Arg Gly Pro Lys Pro Val Trp Gly Leu Pro Gln
120          320          325          330
122 ctg tgc cgg gca ggg ggc ctg ccc cgc cat cct gac gcc gcc ctc ttc 1360
123 Leu Cys Arg Ala Gly Gly Leu Pro Arg His Pro Asp Ala Ala Leu Phe
124          335          340          345
126 ttc cct cct ctg cgc cgc ctc atc ctc ttc aag ggt gcc cgc tac tac 1408
127 Phe Pro Pro Leu Arg Arg Leu Ile Leu Phe Lys Gly Ala Arg Tyr Tyr
128          350          355          360

```

## RAW SEQUENCE LISTING

DATE: 09/09/2002

PATENT APPLICATION: US/10/050,216B

TIME: 12:50:08

Input Set : A:\CORRECTEDSEQUENCE LISTING.txt

Output Set: N:\CRF3\09092002\J050216B.raw

```

130 gtg ctg gcc cga ggg gga ctg caa gtg gag ccc tac tac ccc cga agt 1456
131 Val Leu Ala Arg Gly Gly Leu Gln Val Glu Pro Tyr Tyr Pro Arg Ser
132 365 370 375 380
134 ctg cag gac tgg gga ggc atc cct gag gag gtc agc ggc gcc ctg ccg 1504
135 Leu Gln Asp Trp Gly Gly Ile Pro Glu Glu Val Ser Gly Ala Leu Pro
136 385 390 395
138 agg ccc gat ggc tcc atc atc ttc ttc cga gat gac cgc tac tgg cgc 1552
139 Arg Pro Asp Gly Ser Ile Ile Phe Phe Arg Asp Asp Arg Tyr Trp Arg
140 400 405 410
142 ctc gac cag gcc aaa ctg cag gca acc acc tcg ggc cgc tgg gcc acc 1600
143 Leu Asp Gln Ala Lys Leu Gln Ala Thr Thr Ser Gly Arg Trp Ala Thr
144 415 420 425
146 gag ctg ccc tgg atg ggc tgc tgg cat gcc aac tcg ggg agc gcc ctg 1648
147 Glu Leu Pro Trp Met Gly Cys Trp His Ala Asn Ser Gly Ser Ala Leu
148 430 435 440
150 ttc tgaaggcacc tcctcacctc agaaactggt ggtgctctca gggcaaaatc 1701
151 Phe
152 445
154 atgttcccca ccccggggc agaaccctc ttagaagcct ctgagtcctt ctgcagaaga 1761
155 ccgggcagca aagcctccat ctggaagtct gtctgccttt gttccttgaa gaatgcagca 1821
156 ttgtctttgt ctgtcccccac cacatggagg tgggggtggg atcaatctta ggaaaagcaa 1881
157 aaaagggtcc cagatccctt ggccctttcc tccgaggact tctatcctcc ccaggccttt 1941
158 gtttcttcgg ctaaaggtag agttcctttc aagaggtaac agcactggga tccaagcagg 2001
159 gggatgaaaa actcagcaga gaaattcgag accattttgc aagactgtgc ctttctctc 2061
160 aggacccctt ggtcagttc ttgaaaaacg gtgtcatatt tagtcagagg cccaccccc 2121
161 aggaagcatg gatggggatg aaggcacagg cgtctccaac ctcagaggcc ctttgtggg 2181
162 tcaggacaca gagtgggagg gagactgatg caggcctacc agtccttggc ttttgtctg 2241
163 gggctggaat aaagaggtgc cttcagctgg tgggccgaga aaaaaaaaaa aaaaaaaaaa 2301
164 ggcggccgc 2310
166 <210> SEQ ID NO: 2
167 <211> LENGTH: 445
168 <212> TYPE: PRT
169 <213> ORGANISM: Homo sapiens
171 <400> SEQUENCE: 2
172 Met Val Ala Arg Val Gly Leu Leu Leu Arg Ala Leu Gln Leu Leu Leu
173 1 5 10 15
174 Trp Gly His Leu Asp Ala Gln Pro Ala Glu Arg Gly Gly Gln Glu Leu
175 20 25 30
176 Arg Lys Glu Ala Glu Ala Phe Leu Glu Lys Tyr Gly Tyr Leu Asn Glu
177 35 40 45
178 Gln Val Pro Lys Ala Pro Thr Ser Thr Arg Phe Ser Asp Ala Ile Arg
179 50 55 60
180 Ala Phe Gln Trp Val Ser Gln Leu Pro Val Ser Gly Val Leu Asp Arg
181 65 70 75 80
182 Ala Thr Leu Arg Gln Met Thr Arg Pro Arg Cys Gly Val Thr Asp Thr
183 85 90 95
184 Asn Ser Tyr Ala Ala Trp Ala Glu Arg Ile Ser Asp Leu Phe Ala Arg
185 100 105 110
186 His Arg Thr Lys Met Arg Arg Lys Lys Arg Phe Ala Lys Gln Gly Gly

```

## RAW SEQUENCE LISTING

DATE: 09/09/2002

PATENT APPLICATION: US/10/050,216B

TIME: 12:50:08

Input Set : A:\CORRECTEDSEQUENCE LISTING.txt

Output Set: N:\CRF3\09092002\J050216B.raw

```

187          115          120          125
188 Ala Leu Ala His Ala Phe Leu Pro Arg Arg Gly Glu Ala His Phe Asp
189          130          135          140
190 Gln Asp Glu Arg Trp Ser Leu Ser Arg Arg Arg Gly Arg Asn Leu Phe
191 145          150          155          160
192 Val Val Leu Ala His Glu Ile Gly His Thr Leu Gly Leu Thr His Ser
193          165          170          175
194 Pro Ala Pro Arg Ala Leu Met Ala Pro Tyr Tyr Lys Arg Leu Gly Arg
195          180          185          190
196 Asp Ala Leu Leu Ser Trp Asp Asp Val Leu Ala Val Gln Ser Leu Tyr
197          195          200          205
198 Gly Lys Pro Leu Gly Gly Ser Val Ala Val Gln Leu Pro Gly Lys Leu
199          210          215          220
200 Phe Thr Asp Phe Glu Thr Trp Asp Ser Tyr Ser Pro Gln Gly Arg Arg
201 225          230          235          240
202 Pro Glu Thr Gln Gly Pro Lys Tyr Cys His Ser Ser Phe Asp Ala Ile
203          245          250          255
204 Thr Val Asp Arg Gln Gln Gln Leu Tyr Ile Phe Lys Gly Ser His Phe
205          260          265          270
206 Trp Glu Val Ala Ala Asp Gly Asn Val Ser Glu Pro Arg Pro Leu Gln
207          275          280          285
208 Glu Arg Trp Val Gly Leu Pro Pro Asn Ile Glu Ala Ala Ala Val Ser
209          290          295          300
210 Leu Asn Asp Gly Asp Phe Tyr Phe Phe Lys Gly Gly Arg Cys Trp Arg
211 305          310          315          320
212 Phe Arg Gly Pro Lys Pro Val Trp Gly Leu Pro Gln Leu Cys Arg Ala
213          325          330          335
214 Gly Gly Leu Pro Arg His Pro Asp Ala Ala Leu Phe Phe Pro Pro Leu
215          340          345          350
216 Arg Arg Leu Ile Leu Phe Lys Gly Ala Arg Tyr Tyr Val Leu Ala Arg
217          355          360          365
218 Gly Gly Leu Gln Val Glu Pro Tyr Tyr Pro Arg Ser Leu Gln Asp Trp
219          370          375          380
220 Gly Gly Ile Pro Glu Glu Val Ser Gly Ala Leu Pro Arg Pro Asp Gly
221 385          390          395          400
222 Ser Ile Ile Phe Phe Arg Asp Asp Arg Tyr Trp Arg Leu Asp Gln Ala
223          405          410          415
224 Lys Leu Gln Ala Thr Thr Ser Gly Arg Trp Ala Thr Glu Leu Pro Trp
225          420          425          430
226 Met Gly Cys Trp His Ala Asn Ser Gly Ser Ala Leu Phe
227          435          440          445
230 <210> SEQ ID NO: 3
231 <211> LENGTH: 1335
232 <212> TYPE: DNA
233 <213> ORGANISM: Homo sapiens
235 <220> FEATURE:
236 <221> NAME/KEY: CDS
237 <222> LOCATION: (1)...(1335)
239 <400> SEQUENCE: 3

```

## RAW SEQUENCE LISTING

DATE: 09/09/2002

PATENT APPLICATION: US/10/050,216B

TIME: 12:50:08

Input Set : A:\CORRECTEDSEQUENCE LISTING.txt

Output Set: N:\CRF3\09092002\J050216B.raw

```

240 atg gtc gcg cgc gtc ggc ctc ctg ctg cgc gcc ctg cag ctg cta ctg 48
241 Met Val Ala Arg Val Gly Leu Leu Leu Arg Ala Leu Gln Leu Leu
242 1 5 10 15
244 tgg ggc cac ctg gac gcc cag ccc gcg gag cgc gga ggc cag gag ctg 96
245 Trp Gly His Leu Asp Ala Gln Pro Ala Glu Arg Gly Gly Gln Glu Leu
246 20 25 30
248 cgc aag gag gcg gag gca ttc cta gag aag tac gga tac ctc aat gaa 144
249 Arg Lys Glu Ala Glu Ala Phe Leu Glu Lys Tyr Gly Tyr Leu Asn Glu
250 35 40 45
252 cag gtc ccc aaa gct ccc acc tcc act cga ttc agc gat gcc atc aga 192
253 Gln Val Pro Lys Ala Pro Thr Ser Thr Arg Phe Ser Asp Ala Ile Arg
254 50 55 60
256 gcg ttt cag tgg gtg tcc cag cta cct gtc agc ggc gtg ttg gac cgc 240
257 Ala Phe Gln Trp Val Ser Gln Leu Pro Val Ser Gly Val Leu Asp Arg
258 65 70 75 80
260 gcc acc ctg cgc cag atg act cgt ccc cgc tgc ggg gtt aca gat acc 288
261 Ala Thr Leu Arg Gln Met Thr Arg Pro Arg Cys Gly Val Thr Asp Thr
262 85 90 95
264 aac agt tat gcg gcc tgg gct gag agg atc agt gac ttg ttt gct aga 336
265 Asn Ser Tyr Ala Ala Trp Ala Glu Arg Ile Ser Asp Leu Phe Ala Arg
266 100 105 110
268 cac cgg acc aaa atg agg cgt aag aaa cgc ttt gca aag caa ggg ggc 384
269 His Arg Thr Lys Met Arg Arg Lys Lys Arg Phe Ala Lys Gln Gly Gly
270 115 120 125
272 gcc ctg gcg cac gcc ttc ctg ccc cgc cgc ggc gaa gcg cac ttc gac 432
273 Ala Leu Ala His Ala Phe Leu Pro Arg Arg Gly Glu Ala His Phe Asp
274 130 135 140
276 caa gat gag cgc tgg tcc ctg agc cgc cgc cgc ggg cgc aac ctg ttc 480
277 Gln Asp Glu Arg Trp Ser Leu Ser Arg Arg Arg Gly Arg Asn Leu Phe
278 145 150 155 160
280 gtg gtg ctg gcg cac gag atc ggt cac acg ctt ggc ctc acc cac tcg 528
281 Val Val Leu Ala His Glu Ile Gly His Thr Leu Gly Leu Thr His Ser
282 165 170 175
284 ccc gcg ccg cgc gcg ctc atg gcg ccc tac tac aag agg ctg ggc cgc 576
285 Pro Ala Pro Arg Ala Leu Met Ala Pro Tyr Tyr Lys Arg Leu Gly Arg
286 180 185 190
288 gac gcg ctg ctc agc tgg gac gac gtg ctg gcc gtg cag agc ctg tat 624
289 Asp Ala Leu Leu Ser Trp Asp Asp Val Leu Ala Val Gln Ser Leu Tyr
290 195 200 205
292 ggg aag ccc cta ggg ggc tca gtg gcc gtc cag ctc cca gga aag ctg 672
293 Gly Lys Pro Leu Gly Gly Ser Val Ala Val Gln Leu Pro Gly Lys Leu
294 210 215 220
296 ttc act gac ttt gag acc tgg gac tcc tac agc ccc caa gga agg cgc 720
297 Phe Thr Asp Phe Glu Thr Trp Asp Ser Tyr Ser Pro Gln Gly Arg Arg
298 225 230 235 240
300 cct gaa acg cag ggc cct aaa tac tgc cac tct tcc ttc gat gcc atc 768
301 Pro Glu Thr Gln Gly Pro Lys Tyr Cys His Ser Ser Phe Asp Ala Ile
302 245 250 255
304 act gta gac agg caa cag caa ctg tac att ttt aaa ggg agc cat ttc 816

```

VERIFICATION SUMMARY

DATE: 09/09/2002

PATENT APPLICATION: US/10/050,216B

TIME: 12:50:09

Input Set : A:\CORRECTEDSEQUENCE LISTING.txt

Output Set: N:\CRF3\09092002\J050216B.raw

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:515 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7